DOCUMENT RESUME

ED 196 862 SP 017 182

AUTHOR Sanford, Julie P.
TITLE Comparison of Heterogeneous and Homogeneous Junior High Classes. R&D Report No. 6108.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Feb 80
CONTRACT OB-NIE-G-80-0116
NOTE 80p.: Parts may not reproduce clearly.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Ability Grouping; Academic Ability; *Classroom Techniques; *Comparative Analysis; *Heterogeneous Grouping; *Homogeneous Grouping; Individual Instruction; Junior High Schools; Student Attitudes; Student Behavior; Student Needs; Student Teacher Relationship; Teacher Behavior; *Teacher Effectiveness; Teaching Methods

ABSTRACT

Many secondary teachers perceive a wide range of studennt ability levels within a class as a constraint on effective teaching and classroom management. This study was conducted to determine which levels of teaching effectiveness, which learning environments, and which classroom management problems were associated with very heterogeneous classes, and to examine the characteristics of these classes as opposed to more homogeneous classes. Data were collected by classroom observation, teacher interviews, and case studies in eleven junior high schools. Five general findings are reported: (1) Increased heterogeneity of students' entering academic levels limits the extent to which teachers can successfully adapt instruction to meet the needs of individual students; (2) Teachers in extremely heterogeneous classes may be less able to meet the needs of individual students; (3) Extreme class heterogeneity is associated with a lessened degree of task engagement and student cooperation; (4) Achievement gains of lower ability students may tend to be lower in heterogeneous classes; and (5) Classroom observation and teacher reports indicate that teachers must use diverse strategies to cope with extremely heterogeneous classes. A warning is issued that extremely heterogeneous classes place extraordinary demands on teachers' time, attention, and classroom management skills. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED196862

SP

Comparison of Heterogeneous and Homogeneous Junior High Classes

Julie P. Sanford

R&D Center for Teacher Education
The University of Texas at Austin

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

(R&D Rep. No. 6108)

Research and Development Center for Teacher Education
The University of Texas at Austin

February, 1980

This study was supported in part by The National Institute of Education, Contract OB-NIE-G-80-0116, Classroom Organization and Effective Teaching Project, Carolyn M. Evertson, Project Director, The Research and Development Center for Teacher Education, The University of Texas at Austin. The opinions expressed herein do not necessarily reflect the position or policy of The National Institute of Education and no official endorsement by that office should be inferred.

SP 017 182

Acknowledgements

The conceptualization, planning, and execution of a major study such as the Junior High Classroom Organization Study requires the work and commitment of many people. Before citing individuals, we wish to extend our gratitude to two organizations whose support made this work possible: The Research and Development Center for Teacher Education, Oliver H. Bown, Director, and the Austin Independent School District. School district personnel who assisted us in many ways were Freda M. Holley, Coordinator of Research and Evaluation; Lawrence Buford, Director of Secondary Education; James Patterson and Maud Sims, Assistant Directors of Secondary Education; Margaret Ruska, Language Arts Supervisor; and Elgin Schilhab, Mathematics Supervisor.

Program staff members who made contributions to the design and completion of the study were Barbara Clements, who trained observers and coordinated data collection, and Betty Ancheta, who prepared materials and organized staff participation. The following people were responsible for data collection durirg the full school year: Barbara Clements, Alice Haynes, Nadene Hudson, Julie Sanford, and Patti Shields. They were assisted during the first three weeks of the study be the following observers: Chris Baker, Jane Bowles, Phyllis Brown, Vicki Calderola, David Campbell, Joan Dodds, Susan Guinn, Dean Johnston, Matthew Lee, and Eddie Orum. Data analyses were performed by Jeanne Martin, Donald Veldman, Betsy Galligan, and Mike Kerker. Barbara Clements, Murray Worsham, and Julie Sanford were involved in report preparation. Randall Hickman and Ellen Williams also provided help during the data analysis and reduction stages.

DEC 3 0 1980

The important task of organizing and checking data was performed by Aimee Brodeur, Helen Ross, Sandi Medeiros, and Patty Martinez. Narrative typing was done by Rosemary Brant, Cheshire Calhoun, Candace Grigsby, Diana Hadley, Randall Hickman, and Kathy Woolum. Manuscript and materials preparation was also done by Betty Ancheta, Carol Culp, Susan Smith, and Sheila Haber-Garsombke.

Most importantly, we also wish to thank the principals of the eleven junior high schools and the 51 teachers who allowed us to learn from them.

Comparison of Heterogeneous and Homogeneous Junior High Classes

The issue of ability grouping in schools has been the subject of continuing debate and research. Most of the research and rhetoric has in fact focused on only one aspect of the issue: whether homogeneous ability grouping has a beneficial or detrimental effect on students' academic achievement and self-concept (Esposito, 1973; Heathers, 1969; Johnson, 1970; Waltman, 1979).

Very little research, in contrast, has focused on the learning environment afforded by heterogeneous grouping. Many secondary teachers perceive wide range of students' ability within classes as a constraint on effective teaching and classroom management. In some school districts, desegregation efforts or mainstreaming of learning-handicapped students has put increased pressure on teachers to teach groups of students with very diverse needs and abilities. The purpose of this study was to find out what levels of teaching effectiveness, what learning environments, and what classroom management problems were associated with very heterogeneous classes included in a large-scale study of junior high classrooms.

The data base used in this study was part of the Junior High Classroom Organization Study (JHCOS), conducted by the Classroom Organization and Effective Teaching Project, Research and Development Center for Teacher Education, The University of Texas at Austin. The JHCOS investigated teaching behaviors and classroom management in 102 seventh- and eighth-grade classes in the 1978-1979 school year, including classes from every junior high school in a large urban district (Evertson, Emmer, & Clements, Note 1). The JHCOS was

markable in the amount and depth of information collected in a large number and variety of classes and schools: high, middle, and low SES schools, schools using ability grouping, others with commitments to heterogeneous classes, some schools with students transported from disparate neighborhoods, and schools affected in varying degrees by efforts to mainstream learning-handicapped students. Abundant data were available to compare characteristics of extremely heterogeneous and less heterogeneous classes.

Background

Thorough reviews and discussion of the literature relating to ability grouping in schools have been presented by writers such as Esposito (1973), Heathers (1969), Johnson (1970), and Wilson and Schmidt (1978). The consensus has been that there is little evidence that the practice of homogeneous ability grouping in schools results in greater achievement for students overall and that there is some evidence that it has deleterious social effects on students in lower ability groups. Johnson (1970), Rosenbaum (1976), and other social psychologists warn of increased teacher expectancy effects and social class stratification associated with homogeneous ability grouping or tracking within schools. Teachers, in contrast, generally agree that the teaching and learning environment is maximized by homogeneous grouping (Goldberg, Passow, & Justman, 1966; Wilson & Schmidt, 1978).

Some recent research lends support to the teachers' point of view. Venezky and Winfield (Note 2) reported after a study of schools that demonstrated outstanding success in teaching reading that the successful schools strongly emphasized homogeneous grouping for reading. The authors found range of students' abilities to be a factor limiting the

extent to which teachers can successfully adapt instruction to meet the needs of individual students, especially when classroom support personnel or tutors were not available.

Stallings (Note 3) reported congruous findings from a large-scale study of remedial reading programs in secondary schools in California. Because all of the classes in the study were remedial and all therefore included some students with low reading pretest scores, Stallings defined heterogeneous classes as any classes in which one-fourth to one-half of the students scored over the 49th percentile on the pretest. When classes in the study were subsequently grouped according to achievement gains made, the highest-gain group included no heterogeneous classes. When students' scores within the heterogeneous classes were analyzed, the students who scored over the 49th percentile on the pretest made less gain than did students scoring below the 49th percentile on the pretest. While Stallings (Note 3) recognized that these differences might have resulted from a ceiling effect on the test, she concluded:

> Students in heterogeneous grouped classrooms spent more time making social comments and humorous statements, and spent less time on reading. Thus, our findings suggest that for greater efficiency in teaching and learning reading skills, students should be placed in classes with students of similar reading ability. (p. 65)

A number of additional studies have implications for the ability grouping issue in all subject areas. Metz (1978) reported that in her in-depth study of two schools with tracking systems, teachers used very different classroom routines, management techniques, and teaching strategies with different ability groups of students. This finding contradicted previous studies showing that teachers did not teach

differently in ability grouped classes (Johnson, 1970), but it is consonant with recent findings for teaching effectiveness in elementary schools discussed by Good (1979). Good cited several studies with clear indications that different teaching styles and strategies have different effects on students of different aptitude and self-concept levels. Certain types of students, for example, may benefit more from direct instruction than do other students. Some students need more structure than others. The problems implied in providing optimum instruction for all students in an extremely heterogeneous class are obvious. Good (1979) suggested:

> This does not necessarily suggest that teachers cannot work with the whole class profitably. But such findings do seem to imply that teachers, after a whole class presentation, may need to spend more time with certain students in private interactions. Teachers may want to schedule an extra 10 minutes of instruction for these students later in the day to provide them with additional structure through review, extra explanations and additional practice. (p. 60)

Good's suggestion seems realistic for teachers in self-contained elementary classrooms, but it illustrates the challenge which secondary teachers of heterogeneous classes face when they have only 55 minutes a day with each class and teach 125 or more students.

Methods

Data Sources

The research design and procedures used in the overall JHCOS have been previously presented (Evertson, Emmer, & Clements, Note 1). Only those aspects of overall methodology which relate directly to the present study will be summarized here. A total of 102 seventh- and eighth-grade classes (52 math and 50 English) were observed an average of 14 times each during the year, about six times within the first three

weeks of school and generally one or two hours a month throughout the year. At the end of the year students took math and English achievement tests and completed ratings of the teacher. Also at the end of the year teachers were interviewed and asked to complete a questionnaire, and observers were asked to complete 303 ratings of teachers whom they had seen regularly throughout the year. Students completed ratings of their teachers. Entering California Achievement Test (CAT) scores for all students were obtained from district offices.

Throughout the course of the study, a wide variety of observational data, counts and ratings, student outcome measures, and teacher self-reports were collected for each class. The data which were used in the analysis for this report included:

a) Narratives describing classroom environments, instructional activities, and teachers' and students' behaviors during each observed class session;

b) Component Ratings completed by observers after each observed class session. These consisted of 44 Likert-scale items describing teachers' and students' behaviors and classroom climate;

c) Student Engagement Ratings (SERs) or counts of students in different engagement categories (On-task; On-task, Academic; Off-task, Unsanctioned; or in Dead Time) made every 15 minutes during observations;

d) Students' entering achievement scores;

e) Students end-of-year achievement scores;

f) Student Ratings of the Teacher;

g) 15 teacher self-report items relating to the heterogeneity issue, drawn from the end-of-year teacher questionnaire; and

h) A Management Factor Score for each teacher, derived from factor analysis of 100 reliable items taken from the 303-item end-of-year Observer Rating of the Teacher (ORT).

Subsample Population

Initial inspection of within-class ranges of students' achievement levels as indicated by the standard deviation of entering achievement test scores in each of the 102 classes in the study indicated that many of the more homogeneous classes in the sample tended to be (not surprisingly) "special" classes of all lower achieving or all higher achieving students. This situation presented a potential problem for this study since it was assumed that homogeneous high-ability classes might have very different characteristics than homogeneous low-ability classes. The focus of this study was on characteristics of very heterogeneous classes as oppcsed to less heterogeneous classes. To avoid confounding results with characteristics of high-ability versus low-ability groups of students, it was decided to use only those classes which had mean entering achievement test scores in the 33rd to 67th percentile range. This left us with 27 middle-ability English classes and 20 middle-ability math classes with standard deviations (sigma values) of students' pretest scores ranging from 4.19 to 17.73. Translated into terms of entering grade level equivalents, the least heterogeneous classes in the subsample had ranges of three grade level equivalents, while the most heterogeneous classes had ranges of up to ten grade level equivalents.

It was assumed that because class heterogeneity might be associated with different problems and different solutions in English classes than in mathematics classes, classes in the two subject areas should be

considered separately. There were fewer middle ability math classes (20) than English classes (27) in our subsample. In addition, even after we had narrowed our sample to include only classes that had mean entering achievement test scores in the 33rd to 67th percentile range, there remained a correlation (r = .49) between entering academic level and sigma in math classes. No correlation (r = .08) existed in the group of 27 English classes. Therefore, it was determined that the present report of effects of extreme heterogeneity of students' academic levels within classes should focus on the English classes in our subsample. Table 1 shows data for these 27 classes. The extremely heterogeneous English classes included students with entering reading achievement levels ranging from late first-grade or early second-grade to 11th- or 12th- grade equivalents. The more homogeneous classes were characterized by grade level equivalent ranges of three to four years within individual classes.

Data Analysis

It was hypothesized that many of the classroom variables that might be related to wide range of academic levels of students within a class would be mediated by and partially dependent on the teacher's level of classroom management skills. It was also assumed that even within the somewhat restricted range of class mean entering achievement levels that we used, some difference in classroom climate might be related to differences in the mean student academic level in different classes. We, therefore, chose to utilize a series of regression models that allowed us to control for differences in teachers' management skills and class mean academic levels. In other words, our procedures allowed us to identify those aspects of classroom climate, behavior, and

instructional organization that appear to be affected by wide range of students' entering ability, above and beyond differences that could be predicted from knowledge of the teacher's management skill level or general student academic level in a classroom.

It was also hypothesized that teachers who were effective classroom managers might respond differently to, or deal differently with, extremely heterogeneous classes than would teachers who were poorer classroom managers. Our models were, therefore, designed to allow us to test for interactions between class heterogeneity and teachers' management skill level.

Following is the series of regression models which were evaluated for 27 English classes:

1. Criteria = U + CAT78 + MFS + Sigma + Sigma*MFS + Error
2. Criteria = U + CAT78 + MFS + Sigma + Error
3. Criteria = U + CAT78 + MFS + Error

Where:

U = a standard unit vector used in computation.

CAT78 = class mean total raw score on the entering California Achievement Test in reading.

MFS = the Management Factor Score for each teacher, which was derived from factor analysis of the end-of-year Observer Ratings of Teacher. Individual ratings contributing to the management factor covered all aspects of classroom management in both of the teacher's two classes. The Management Factor Score served as a general index of the teachers' classroom management skill level.

Sigma = the standard deviation of entering achievement scores (CAT78) within classes. Sigma served as the index of class heterogeneity or range of entering achievement levels within classes.

Sigma*MFS = the interaction of class heterogeneity and the teacher's Management Factor Score.

Criteria. The criteria used in the evaluation of the above models included:

1. Mean class scores on 44 Component Ratings describing different aspects of classroom environment, instruction, student behaviors, and classroom management strategies;

2. Class means (percentage of students) on four measures of student task engagement: On-task; On-task, Academic; Off-task, Unsanctioned; and in Dead time;

3. Teachers' responses to 15 questions relating to the class heterogeneity issue, drawn from a teacher questionnaire.

4. Class mean achievement gains indicated by end-of-year achievement test scores; and

5. Class mean Student Ratings of the Teacher (end-of-year).

Evaluation of the models proceeded in the following manner. Model 1 (full model) was compared to Model 2 to test the significance of the interaction of sigma and teachers' Management Factor Scores in the prediction of the criterion variable. Subsequent plotting of predicted values associated with significant ($p \leq .05$) interactions indicated the direction and magnitude of the interrelationships of teachers' management skill level, class heterogeneity, and the particular criterion variable.

Model 2 was compared to Model 3 to test the significance of class range of students' entering achievement levels (sigma) in predicting criteria. In other words, this comparison evaluated the extent to which knowledge of classroom heterogeneity allowed us to predict a classroom variable better than with knowledge of only class mean entering achievement levels and teachers' classroom management skills.

One additional hypothesis contributed to our plans for evaluating the regression models. It seemed likely that extreme heterogeneity of students' academic levels within classes might be associated with more apparent problems at the beginning of the school year than later in the year, when teachers have more knowledge of their students. Therefore, it was decided that analyses would be done separately for Component Ratings and Student Engagement Ratings data taken during observations in the first three weeks of school (about six observations per class) and in the rest of the year (about eight observations per class).

Within-class Analysis

The regression models explained above would give us little information about individual students or different types of students in extremely heterogeneous classes. In order to get some idea of possible effects of extreme heterogeneous ability grouping on achievement of individual students of different academic levels, students were categorized in low (below 33rd percentile), middle (33rd to 67th percentile), and high (above 67th percentile) entering achievement levels in each of the 27 classes in the study. Students in each class were also categorized according to their achievement gains during the year of the study: low gain (below -.25 standardized residual gain), medium or expected gain (residuals from -.25 to +.25), and high gain

(residuals above +.25). The resulting 3 x 3 plot for each class showed whether the teacher was equally successful with all three ability groups in the class.

The percentage of students with expected (medium) or above expected achievement gains in each category of entering achievement level were computed in each class. Correlation coefficients between each percentage figure and sigma were obtained. This procedure gave some idea of whether, as class heterogeneity increased, achievement gains among lower-ability students (or higher-ability students) tended to suffer (or increase).

Case Studies

Finally, the most heterogeneous English classes in our subsample were selected for descriptive case studies of narratives taken during classroom observation. All of the classroom narratives for five classes were read for information relating to heterogeneity of students' academic levels in the class. All of the case study classes had students of entering achievement levels spanning eight or more grade level equivalents. Some of the classes were taught by very experienced and effective classroom managers. Others were taught by less effective classroom managers. Our narrative study focused on the organizational and instructional strategies teachers used (or attempted to use) to accommodate the diverse needs and abilities of their students. The end-of-year interviews with these five teachers were also read for additional information.

Results

Results of the above described analyses suggest that extreme heterogeneity of students' academic levels within English classes

significantly affects a number of aspects of classroom climate and teaching effectiveness. The regression models for 27 English classes were evaluated with 113 tests (48 classroom variables for two periods of the school year plus 17 end-of-year measures). These tests resulted in 22 main effects significant at the $p \leq .05$ level, and 11 significant interactions ($p \leq .05$), far exceeding the number of findings that would be expected from chance.

Component Ratings

Component Ratings made after each classroom observation resulted in mean scores on 44 measures of classroom climate, instructional organization, and behavior management in each class. Regression analysis indicated that class heterogeneity was strongly related to two blocks of Component Rating variables: those describing positive affect or interpersonal teacher-student relationships, and those relating to academic fit, or observers' impressions of the extent to which lessons were suited to the ability levels, interests, and backgrounds of the students.

Table 2 lists the Component Rating variables which relate to academic fit, and shows the results of the tests for relationships between these variables and class heterogeneity. Several conclusions may be drawn from the findings shown.

First, low mean scores (1.78 to 1.94 on a 5-point scale) for two variables, Assignments for different students and Variety of materials, for both periods of the school year indicated that on the whole, teachers in our sample did not use a lot of differentiation of assignments or materials for individuals or small groups. They tended to use whole-class activities and a basic set of materials.

Furthermore, observer ratings of Different assignments and Variety of materials did not tend to be significantly higher in very heterogeneous classes. Providing different assignments or increased variety of materials for individuals or small groups was not a strategy the English teachers used to cope with wide range of student ability in their classes overall. There was a significant interaction effect later in the school year for Variety of materials. This will be discussed later in this section.

Second, at the beginning of the school year, both good and poor classroom managers tended to have more trouble accommodating their lessons to all of their students in very heterogeneous classes, than in more homogeneous classes. During the first three weeks of school, extreme class heterogeneity was associated with lower observer ratings on:

Degree of student success ($p \leq .01$);

Content related to students' interests and backgrounds ($p \leq .05$);

Lessons adapted to different student ability levels ($p \leq .05$); and

Materials effectively support instruction ($p \leq .05$).

These findings seem reasonable considering the teachers' lack of knowledge of his/her students at the beginning of the year, especially in the face of an eight to ten year grade equivalent spread in the more heterogenous classes. No heterogeneity effects significant at $p \leq .05$ were found for seven other academic fit variables in the first three weeks.

Third, after the first three weeks of school, effective classroom managers in English classes were able to adjust to diverse academic needs of their students in very heterogenous classes, but less effective

classroom managers were not able to do so. The main effects or significant relationships which showed up in the first three weeks of school disappeared in the "rest of year" analysis. Not a single academic fit variable was significantly related to class heterogeneity after the first three weeks when main, or overall, effects were evaluated. However, significant interaction effects appeared, indicating in every instance that teachers who had high Management Factor Scores were able to accommodate the diverse academic needs of their students in heterogenous classes, but poor classroom managers were not able to do so.

Figures 1 through 5 are plots of the significant interactions obtained for the following variables:

Students' attention spans are considered (Figure 1);

Variety of materials (Figure 2);

Presentation is clear (Figure 3);

Lesson is adapted to different academic levels of students (Figure 4);

High degree of student success (Figure 5).

These interaction plots should be interpreted in the following manner: The X-axis includes only two points, one representing classes with relatively low standard deviations (sigma values) of students' entering academic levels, and one representing classes with high sigma values (i.e., extremely heterogenous classes). The two lines on the graph connect points indicating the predicted Component Ratings of two different groups of teachers: those with relatively high Management Factor Scores (HI MFS) and those with relatively low Management Factor Scores (LO MFS). For example, in Figure 1 the plotted points indicate

that in relatively homogeneous classes (low sigma), observers were likely to rate both good classroom managers and poor classroom managers fairly high on presenting lessons which take the attention spans of students in the class into consideration. In contrast, in very heterogenous (high sigma) classes, observers were likely to notice problems with students' attention spans in classes of poor managers, thus the very low value for the point, marked "LO MFS, HI Sigma." (Since the points represent predicted values, rather than actual scores, it is possible to have values below 1, although the Component Ratings were on a 1 to 5 scale.) Teachers with high management factor scores, in contrast, were apparently able to accommodate diverse attention spans in heterogenous classes, at least to the extent that observers did not notice problems which would cause them to give much lower ratings to these teachers in extremely heterogenous classes.

Interactions depicted in the other graphs should be interpreted in similar fashion. Taken together, they indicate that the effective classroom managers in our sample of English teachers made realistic and effective adjustments to extremely heterogenous class situations after the first three weeks of school, but poor classroom managers did not.

Another block of Component Rating variables related to class heterogeneity was a set of classroom climate variables describing positive affect or personal student-teacher interactions in the classroom. Table 3 shows these nine variables and results of regression analysis for each in both periods of the school year. Findings for variables in the first three weeks of school were limited but clear. Teachers in extremely heterogenous classes were rated by observers as being less oriented to student personal needs, less receptive to student

input, and devoting less class time to nurturance of affective skills. For the majority of the school year, significant negative relationships were seen between class heterogeneity and seven of the nine positive affect variables:

Teacher states desired attitudes ($p \leq .05$);

Amount of positive reinforcement ($p \leq .001$);

Teacher's listening skills ($p \leq .01$);

Teacher expresses feelings ($p \leq .001$);

Teacher is receptive to student input ($p \leq .01$);

Teacher is oriented to student needs ($p \leq .001$); and

Nurturance of affective skills ($p \leq .05$).

One significant interaction is shown in Figure 6. This finding could be interpreted several ways, but in general it does not conflict with the main effect indications that teachers were able to devote less time and energy to affective concerns of students in extremely heterogeneous classes.

The remaining Component Rating variables fall into two categories: those describing instructional management variables not necessarily related to the academic fit problem; and those describing class behavior management variables. Table 4 shows the results for instructional management variables. This category overlapped in some cases with instructional or academic fit variables discussed earlier, and three of these variables were included in both tables. With the exception of the three overlap variables (Variety of materials, Materials effectively support instruction, and Presentation is clear), only one variable in this category (Teacher provides and seeks rationale/analysis) was significantly related ($p \leq .05$) to spread of academic levels of students

in the class. Observers noted less discussion focusing on rationale and analysis in very heterogenous classes in the first three weeks of school. This finding might be related to the trend suggested by positive affect variables. Perhaps teachers in extremely heterogenous classes have less time for affective or higher cognitive discussions in general. No significant relationships were found for the rationale and analysis variable after the first three weeks, however. In general, then, teachers were not found to vary their instructional management or lesson presentation style with increasing class heterogeneity.

There also appeared to be relatively little relationship between class heterogeneity and 20 classroom behavior management variables (see Table 5). There were four findings significant at $p \geq .05$. Observer ratings indicated that in very heterogenous classes, disruptive behavior, when it occurred, came from a greater number or variety of students. There were not, however, greater amounts of disruptive behavior rated in heterogenous classes than in more homogeneous classes. For the majority of the school year, amount of positive reinforcement was strongly related ($p \leq .001$) to class heterogeneity. Observers rated very heterogenous classes lower in amounts of positive reinforcement used by the teacher. This variable was also included in the set of variables denoting positive affect, since it related to classroom climate and student self-concept, as well as to teachers' behavior management strategies.

There was also a negative relationship between ratings of teachers' consistency in dealing with behavior and class heterogeneity. In view of previous research findings (Metz, 1978) that different management strategies work with different types of students, it seems quite likely

that in very heterogenous classes, teachers may have to use behavior management strategies that would appear inconsistent to an observer (and possibly to some students themselves).

Finally, Figure 7 shows an interaction between class heterogeneity and teacher's management skill level, with respect to the variable, Teacher desists inappropriate behavior (i.e., the teacher tries to stop inappropriate student behavior without personally criticizing student). Poor classroom managers used this behavior management strategy much less than did effective classroom managers in relatively homogeneous classes. But in extremely heterogenous classes, observers rated poor classroom managers higher on such desists, while good classroom managers were rated as using them less. One interpretation is that poor classroom managers ignored inappropriate behavior which stayed at tolerable levels in homogeneous classes, but that in extremely heterogenous classes, the levels or kinds of inappropriate behavior in classes of poor managers tended to be such that teachers could not ignore them and still carry on the class, thus they were forced to constantly use desists. This is only one interpretation, but it is supported by results obtained for Student Engagement Ratings.

Student Engagement Ratings

Student Engagement Ratings (SERs) were more sensitive measures than Component Ratings, because they were low-inference counts of students in different engagement categories, taken every 15 minutes during observations. Table 6 shows results of analysis of the relationships between class heterogeneity and classroom percentages of students Off-task, Unsanctioned; On-task, Academic; On-task; and in Dead Time. There were no significant relationships ($p \leq .05$) between class range of

students' academic level and percent of students in dead time, or percent of students on-task, academic. Throughout the school year, however, there were lower percentages of students on task (doing what they were supposed to be doing, either academically or procedurally) in extremely heterogenous classes. During most of the year, but not in the first three weeks of the year, there were also higher percentages of students counted as off-task, unsanctioned in more heterogenous classes ($p \leq .01$).

In addition to the main effects, three interactions between classroom heterogeneity and Student Engagement Ratings were significant at $p \leq .05$. These interactions show a pattern very consistent with that of the academic fit variables in English classes. Student behavior variables in classes of teachers who were generally good classroom managers were slightly negatively affected by increasing class heterogeneity, but student behavior variables in classes of poor classroom managers appeared to be markedly affected by extreme class heterogeneity.

Figure 8 shows that Off-task, Unsanctioned percentages were very low in relatively homogeneous classes taught by effective classroom managers and increased only slightly with increasing heterogeneity of students' academic levels in classes. Off-task percentage rates for classes taught by poor classroom managers, in contrast, were much higher in extremely heterogenous classes than they were in relatively homogeneous classes. Figure 9 shows corresponding patterns for percentage of students on task. Students' on-task rates decreased slightly with increased heterogeneity in classes taught by effective classroom managers, but decreased markedly in classes taught by teachers

with low Management Factor Scores. Figure 10 shows that in relatively homogeneous classes, poor classroom managers had high percentages of students on-task, academic; but in extremely heterogenous classes, they had much lower academic engagement rates. Good classroom managers, in contrast, actually had higher percentages of students academically on-task in very heterogenous classes than in more homogeneous classes.

Taken together, the main effects and interactions resulting from regression analysis of Student Engagement Ratings in English classes indicate that extreme class heterogeneity is associated with a lessened degree of student cooperation and student task engagement in classes overall, and particularly in classes taught by relatively poor classroom managers.

Teacher Questionnaire Items

The questionnaire that teachers completed at the end of the Junior High Classroom Organization Study included 15 self-report items pertaining to the ability-grouping issue. These items are listed in Table 7. They include three estimates of class time use, eight items describing frequency of use (on a 1 to 3 scale) of different strategies to cope with varying ability levels in class, one question about grading emphasis in mixed-ability classes, and three items describing related teacher concerns. The last four items had response scales of one to five.

First, on a descriptive level, the questionnaire items suggest a number of things about our subsample of English classes. As a group, the teachers reported spending most of their class time (mean response of 62%) in whole-class instruction. This self-report was corroborated by classroom observation data. In response to the question about how

they coped with a wide range of student ability in their classes, the most highly rated responses were: differential grading and testing, ability grouping within classes, and use of special attention or conferences with individual students. Teachers reported a low mean frequency of use of student teachers, aides or resource persons to help cope with varying ability levels of students in class. Teachers reported a fairly high emphasis on student effort when they assigned grades in mixed-ability classes and high mean levels of concern with meeting needs of different students. The mean response to the question about the extent to which academic range of students was a problem or constraint was 2.52, but responses ranged from 1 (not a problem) to 5.

Teachers' responses to the 15 questionnaire items were analyzed with our set of regression models to evaluate the extent to which responses on any of the items were related to within-class range of student academic levels. The analysis resulted in two main effects and one interaction with teacher's classroom management skill level. Teachers of extremely heterogeneous classes tended to report higher frequency of use of ability grouping within class ($p \leq .05$) and peer tutoring ($p \leq .01$). Teaching extremely heterogeneous classes did not appear to increase the probability of teachers reporting greater use of any of the other instructional strategies or expressing greater concern for meeting the needs of different students.

An interaction obtained for the self-report variable, Percentage of class time devoted to individuals in class (Figure 11), indicated that teachers who had high management factor scores reported devoting more time to individual students in extremely heterogeneous classes. In contrast, poor classroom managers reported high percentages of time

spent with individual students in relatively homogeneous classes, but relatively little time devoted to individuals in very heterogeneous classes. This interaction finding is interesting in that, although it is based on a teacher self-report rather than on classroom observation, it closely mirrors the pattern seen in interactions obtained with Component Ratings related to academic fit problems in heterogeneous classes. (For example, compare Figure 11 with Figure 2, Variety of Materials.) In general, teachers who were poor classroom managers appeared to respond ineffectively to the demands of extremely heterogeneous classes, while teachers with good overall classroom management skills made modest, but at least to some extent, effective accommodations to the diverse needs of students in extremely heterogeneous classes.

Outcome Variables

Outcome variables used in the overall JHCOS were class mean adjusted achievement gains and Student Ratings of the Teacher (SRT). Table 8 shows that neither of these measures was significantly related to range of students entering academic levels in classes. Considering the consistent, significant relationships between class heterogeneity and positive affect or teacher-student interaction variables shown in Table 3, a relationship between classroom heterogeneity and students' attitudes toward the teacher might be expected. If it existed, it remained unmeasured by our Student Rating of the Teacher instrument.

Findings for significant relationships between class heterogeneity and instructional fit variables (Table 2) and between class heterogeneity and Student Engagement Ratings (Table 6) would in a similar fashion appear to point to possible effects for achievement

gain. But no significant, or near significant, relationship was found between class heterogeneity and mean class achievement gains, using our regression models that controlled for entering mean class achievement scores and teachers' classroom management skill levels. However, the mean class achievement gain figures told us almost nothing about achievement gains of different ability groups of students within very heterogenous classes. To answer that question, we turned to the within-class analysis procedure described in the Methods section of this report.

Within-class Analysis

Within-class analysis of gains made by lower and higher academic level students in English classes yielded the correlations shown in Table 9. Students in each class were categorized into high, medium, and low entering achievement groups and into higher-than expected, expected, or below-expected achievement gain categories. The percentage of students in each entering achievement group who made expected or above-expected achievement gains was computed for each class, and these percentages were correlated with sigma, or class heterogeneity. This correlation for the high entering achievement group did not approach significance. Extreme class heterogeneity was not associated with lower gains for the higher entering achievement level students in each class. In other words, extreme heterogeneous grouping did not appear to handicap higher academic level students in those classes.

There was a significant negative correlation ($R = -.33$, $p \leq .05$) between percentage of low entering achievement students with expected or above expected gains and class heterogeneity. This finding suggests that extreme heterogeneity may be a handicap to academic progress of the

lower ability students in English classes. This does not indicate that lower academic level students achieve better in extremely homogeneous low ability groupings; no very homogeneous remedial level classes were included in our subsample, but lower ability students in our study tended to achieve better in moderately mixed (or moderately homogeneous) classes than in extremely heterogeneous classes in which teachers had to adjust instruction to student academic levels spanning equivalents of eight or more school years.

Case Studies

All the classroom narratives for five of the most heterogeneous classes in our sample of English classes were read for additional insights into class climate, instructional organization, and classroom management in classes having students of widely varying academic levels. Descriptive data for these classes are shown in Table 10. The case study findings for each class will be briefly summarized in this section of the report, then trends seen in the five classes will be discussed.

Class A included students whose entering achievement levels spanned nine grade level equivalents. The class mean achievement gain was relatively high. The teacher was rated by observers as an effective classroom manager and was well-liked by her students. For the most part, Teacher A used whole-class activities in her seventh-grade class. But several strategies for meeting different student needs and abilities were evident. First, she consistently gave extra attention and assistance to the lower level students, both during regular class time and before and after school. She arranged students' desks so that students who needed the most help were seated near her desk. She frequently checked with these students at the beginning of seatwork

activities to make sure they understood directions and to give them extra assistance if necessary.

Teacher A provided all of her students with very frequent academic feedback. Many of her lessons featured several short seatwork assignments or practice activities, which were checked and discussed during class. The teacher appeared to pay careful attention to how individual students were doing. She kept all of the students responsible for participation in class discussion or recitation by systematically including everyone in these activities, calling on volunteers and nonvolunteers to participate.

Students were seldom grouped for instruction, even for spelling. In her interview, the teacher commented that in previous years she had provided individual contract work for low academic level students, but found that students became bored with that approach. During the year of the study, students in Class A all did the same assignments, but the teacher sometimes used differential grading, reducing the portion of assignments which some students were expected to complete or, rarely, excusing them unobtrusively from certain tests or activities and allowing them to make up the grade with a substitute assignment later. Occasionally, she directed one student to help another in class on a particular task. Teacher A did not emphasize extra credit work or provide enrichment assignments for the higher ability students in the class. Some students, in fact, were frequently noted in dead time, waiting quietly for the next activity or for the end of the class. Nevertheless, within-class analysis of residual achievement gains indicated that only one of the higher-ability students in Class A made low achievement gains.

Class B was an eighth-grade class with students' entering achievement scores ranging from Grade Level 2.4 to 11.4. Mean achievement gains in the class were close to mean gains for all of the middle-ability English classes. Student Ratings of the Teacher were rather low, although observers rated the teacher as an effective classroom manager. Like Teacher A, Teacher B gave all of her students frequent academic feedback and actively encouraged a high level of student responsibility for their work. She did not devote much class time to providing individual assistance to lower-ability students, however. Instead, she repeatedly encouraged students who needed more help or more time to complete assignments to come in after or before school. We have no information about whether students did so. Students were grouped for spelling and had different assignments. Observers did not record use of grouping in other content areas, provision of much extra credit work or enrichment materials, or peer tutoring. Little Dead Time was noted in Class B.

Class C had the lowest mean entering achievement level and the highest sigma value of our five case study classes. The teacher was rated low by both students and observers. Students, nevertheless, achieved at expected levels, with achievement gains evenly distributed across low, middle, and higher entering achievement groups of students within the class. Narratives showed very little evidence that the teacher made any adjustment to accommodate different student needs and abilities. Students were grouped for spelling, but not for other activities. Little or no provision of extra credit or enrichment materials, peer tutoring, or additional attention or assistance for slower students was noted in the narratives. Many lessons featured

whole-class recitations and drill for long periods of time. General student confusion and boredom was often noted by classroom observers, but the teacher generally kept students academically engaged in class. Some students characteristically finished their work early. Others never seemed to finish. In the classroom observations and in the teacher interview, Teacher C evidenced a low level of awareness of individual students. When questioned in her interview, she said the lower level students in this class (unlike those in classes she had taught in previous years) wanted to work on the same level as the rest of the class and refused to do lower level work at which they could succeed. She attributed a number of problems in her class to heterogeneous grouping practices and to the general character of students in her classes.

Class D was in the same school as Class C. The teacher was rated highly by classroom observers (with respect to the Management Factor Score) and by students. Mean achievement gains for the class were somewhat low, however. The higher-ability students in this class made better gains as a group than either the middle- or lower-ability students. This teacher used grouping within class more than did the teachers in the other case studies. In addition to having three or four spelling groups, the teacher often used flexible or informal grouping of students for some other activities, such as handwriting and composition critique sessions. High academic level students were provided with a file drawer of enrichment materials and activities, which were frequently changed. Lower students got some degree of extra assistance and attention from the teacher in class. Poor readers were provided with optional audiotapes and headphones when students read silently in

class. Frequent feedback, high levels of student accountability for their work, and little Dead Time in class were noted.

Teacher E was rated most highly by her students and had a higher-than-mean Management Factor Score. Mean gains in the class were low, with middle and higher academic level students making better gains than the lower-ability groups in the class. The class had only whole-class activities, except for some spelling groups in the second half of the year. No grouping was used for other activities, little extra credit or enrichment was provided, and no peer tutoring was noted.

Moderate levels of academic feedback and student accountability for work were described in the narratives for Class E. The teacher sometimes checked with lower level students at the beginning of seatwork to clarify directions. Slower students were very often allowed to finish assignments at home, while faster students often finished early, resulting in some Dead Time for these students. Rather than conferring with slower students individually, the teacher frequently announced publicly that students who had not finished the assignments could finish at home. Many students in the class may have therefore felt little pressure to use their class time productively. In interviews, the teacher indicated that she sometimes applied different grading standards to a few students' work, in order to compensate for different ability levels.

Teachers A through E varied in the extent to which they recognized and actively tried to accommodate the different levels of students in their classes. In comparing the classes, several trends were noted. First, none of the teachers attacked the problem of extreme class heterogeneity by using innovative or unique patterns of instructional

organization. For the most part, they used a whole-class approach. The more effective teachers (based on observer ratings and student attitudes as well as on achievement gains) accommodated different students by supplementing whole-class activities with additional in-class attention and assistance for lower-level students. They also used some grouping and different assignments, at least in spelling; and they occasionally used peer tutoring. No clear patterns were seen in provision of differentiated assignments or enrichment materials for the higher-ability students in the class. In every class, some students sometimes finished work early and others did not finish. Teachers dealt with this predictable problem in different ways, but frequent academic feedback to all students and maintenance of high levels of students' responsibility for written work and participation in class for all students appeared to be an important factor in the case study classes.

Summary and Discussion

In this section we will summarize and comment on five general findings of this study of heterogeneous ability grouping in junior high school English classes. First, increased heterogeneity of students' entering academic levels does limit the extent to which teachers can successfully adapt instruction to meet the needs of individual students. Observer ratings of such classroom variables as: considering students' attention spans; adapting lessons to students' ability levels; relating content to students' interests and backgrounds; and high degree of student success were all negatively related to increased class heterogeneity. Such an effect would be expected, and it closely corroborates findings reported by Venezky and Winfield (Note 2). It is important to note, however, that after the first three weeks of school

effective classroom managers were able to make the adjustments necessary to neutralize these negative effects, but less effective classroom managers were not. Students and teachers in some of our very heterogeneous classes, then, continued to suffer the effects of academic fit problems throughout the school year.

Second, there was a strong pattern of findings indicating that teachers in extremely heterogeneous classes may be less able to meet the affective needs of their students. They were rated lower on such variables as: teacher's listening skills; teacher expresses feelings; teacher is receptive to student input; teacher is oriented to student needs; and nurturance of affective skills. Interpretation of findings in this area should be tempered by the fact that in tests of interrater reliabilty, Component Ratings describing positive affect were found to be less reliable than those describing academic fit. Nevertheless, the strong pattern of results for this whole block of variables warrants consideration. A plausible explanation for the trend is that the extra demands that extreme class heterogeneity places on teachers' class time and attention limits their ability to attend to affective concerns of their students. There is simply less time for relaxed conversation, discussion of nonacademic topics, and listening to what students have to say. It was not surprising, then, that teachers in the five case studies of extremely heterogeneous classes were consistently characterized as very businesslike.

Third, Student Engagement Ratings indicated that extreme class heterogeneity is associated with a somewhat lessened degree of task engagement and student cooperation. This effect was much greater in classes taught by relatively poor classroom managers, but there was

evidence of some differences even in classes taught by better classroom managers. Several aspects of heterogeneous classes might contribute to such behavior differences. In a class in which the teacher is less able to adjust instruction to meet individual needs and differences, there are likely to be higher incidences of frustration and boredom among some students. In addition, increased demands on teachers' time and attention make it more difficult to monitor students' behavior constantly. Additionally, perhaps if peer tutoring is used in an effort to cope with different ability levels in a class, opportunities for off-task behavior may be increased. A number of aspects of heterogeneous classes might contribute to student behavior problems.

Fourth, extreme class heterogeneity was not shown to be related to mean class achievement gains or Student Ratings of the Teacher. There was indication, however, that achievement gains of lower ability students may tend to be lower in extremely heterogeneous classes, while achievement gains of middle and high academic level students do not appear to be affected. Study of narratives for five case study classes showed that those teachers directed most of their observable efforts to differentiate assignments and materials, and provide individual attention to the lowest-level students in their heterogeneous classes. In all five classes, however, greater percentages of middle- or high-ability students made good achievement gains than did low-ability students. (In the overall JHCOS, high residual gains did not tend to be made by high-ability students more than by lower-ability students.)

Findings of no significant relationship between class heterogeneity and mean class achievement or mean class SRT corroborate results previously reported for the Texas Junior High School Study (Note 4).

Findings of the within-class analysis of student achievement in heterogeneous classes are compatible with those reported by Stallings (Note 3) and Venezky and Winfield (Note 2) to the extent that they tend to favor more homogeneous groupings. They do not, however, support Stallings' (Note 3) finding that the gain in heterogeneous classes in her sample came more from low-ability students than from high-ability students. One explanation for this apparent discrepancy is that samples in the two studies were really not comparable. Since all the classes in the Stallings' study were basic or remedial junior high reading classes, the classes labeled "heterogeneous" in that study did not contain the extremes of high-ability as well as low-ability classes that the heterogeneous classes in the JHCOS did.

Finally, in answer to the question, "How do English teachers cope with wide ranges of ability in their classes?" classroom observation and teacher self-reports indicate that teachers in our sample used the following strategies with extremely heterogeneous classes:

a) special attention and in-class assistance to lower ability students in the class;

b) limited use of within-class grouping and differentiation of materials or assignments. Except for spelling, for which teachers often grouped students, most teachers only supplemented whole-class activities with remedial (or to a lesser extent) enrichment materials for certain students.

c) some differential grading, taking individual student levels and relative progress into consideration when assigning grades;

d) limited use of peer tutoring; and

e) provision of frequent academic feedback to all students in class and maintenance of high levels of student accountability for both written work and for participation in class.

In conclusion, results of this study do not amount to a caveat against heterogeneous ability grouping in English classes. They do warrant a loud warning: If, in consideration of the social psychologists' case against homogeneous ability grouping and tracking, school systems feel impelled to abandon ability grouping and "special" classes for some students, then they must recognize that the extremely heterogeneous classes that result are indeed also "special." They place extraordinary demands on teachers' time, attention, and classroom management skills. Our study indicates that skilled classroom managers

be less than ideal learning environments. Limiting the number of students in a class or providing classroom aides would probably improve conditions in extremely heterogeneous classes. Finally, findings of this study suggest that school districts should place a high priority on helping their teachers improve their classroom management and organizational skills.

Reference Notes

1. Evertson, C., Emmer, E., & Clements, B. Report of rationale, methodology, and instrumentation of the Junior High Classroom Organization Study (R&D Rep. No. 6100). Austin, Texas: Research and Development Center for Teacher Education, August, 1979.
2. Venezky, R. L., & Winfield, L. F. Schools that succeed beyond expectations in teaching reading. Report to National Institute of Education for Grant NIE-G-78-0027. Newark, Delaware: University of Delaware, 1979.
3. Stallings, J. A. A study of basic reading skills taught in secondary schools. Report of Phase I findings. Menlo Park, California: Stanford Research International, January 1978.
4. Evertson, C., Anderson, L., & Brophy, J. The Texas Junior High School Study: Final report of process-outcome relationships (R&D Rep. No. 4061). Austin: Research and Development Center for Teacher Education, The Univeristy of Texas at Austin, 1978.

References

Esposito, D. Homogeneous and heterogeneous ability grouping: principal findings and implications for evaluating and designing more effective educational environments. Review of Educational Research, 1973, 43, 163-179.

Goldberg, M., Passow, A., & Justman, J. The effects of ability grouping. New York: Teachers College Press, Columbia University, 1966.

Good, T. L. Teacher effectiveness in the elementary school. Journal of Teacher Education, 1979, 30 (2), 52-64.

Heathers, G. Grouping. In R. L. Ebel (Ed.). Encyclopedia of

1970

Metz, M. H. Classrooms and corridors. Berkeley: University of California Press, 1978.

Rosenbaum, J. Making inequality. New York: Wiley, 1976.

Waltman, B. Ability grouping: separate and unequal? Journal of Thought, 1979, 14 (3), 215-220.

Wilson, B. J., & Schmidt, D. W. What's new in ability grouping? Phi Delta Kappan, April 1978, 535-536.

Table 1

Subsample of Average-Ability[1] English Classes

Class Number	Grade	Management Factor Score	Mean Entering Raw Score CAT	Sigma	Grade Level Range	Number of Students
3.5	8	106	49.75	14.45	3.5-11.4	20
3.6	8	106	37.33	11.29	2.4-9.6	18
5.5	7	116	41.84	10.19	2.1-9.1	25
6.5	8	82	48.23	13.89	3.3-11.6	27
7.4	7	114	39.50	14.21	1.9-11.4	20
7.5	7	114	43.43	12.35	4.2-10.5	21
11.3	8	113	46.95	13.41	1.7-10.5	21
11.5	8	113	47.67	14.04	2.4-11.4	24
14.1	7	86	37.82	17.24	2.4-10.9	22
14.2	7	86	40.50	14.75	2.3-12.7	18
16.1	7	118	44.8	17.17	2.4-11.6	22
16.3	7	118	42.37	15.41	2.1-11.4	16
18.2	7	83	39.00	13.92	3.5-11.1	24
20.5	8	81	42.91	9.12	5.3-10.5	23
22.4	8	85	37.58	9.53	4.2-9.1	19
22.6	8	85	39.00	9.95	4.2-8.9	29
25.5	7	105	32.35	6.59	2.6-7.3	26
26.3	7	108	40.29	14.91	2.6-11.4	17
27.1	8	110	50.5	11.7	5.4-13.5	16
27.4	8	110	46.00	11.35	3.9-10.2	22
38.4	8	107	42.68	4.62	5.6-9.1	19

Table 1-continued

Class Number	Grade	Management Factor Score	Mean Entering Raw Score CAT	Sigma	Grade Level Range	Number of Students
45.5	7	88	41.75	10.59	2.6-10.4	28
46.1	7	76	37.74	8.75	3.3-9.1	27
48.2	8	109	49.2	9.03	5.1-10.7	25
48.5	8	109	45.23	8.33	4.6-7.5	22
53.3	8	66	48.45	8.28	5.6-9.8	20
53.6	8	66	48.62	9.37	3.3-9.6	21
Means		98.5	43.02	11.65		
Standard Deviation		16.4	4.70	3.19		
Total Range		66-118	32.35-50.50	4.62-17.17		16-29

[1]Class Mean Entering CAT in 33rd to 66th Percentiles

Table 2

Results for Academic Fit Variables and Class Heterogeneity in English Classes

Variables	First 3 Weeks						Rest of Year					
	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Describes objectives clearly	3.81	.88	.0055	ns	.0191	ns	4.04	.55	.0665	ns	.0086	ns
Attention spans considered	3.41	.88	.0771	ns	.0052	ns	3.44	.84	.0412	ns	.2382	.001
Assignments for different students	1.94	.93	.0109	ns	.0022	ns	1.78	.57	.0545	ns	.0282	ns
Occurrence of verbal class participation	3.14	.98	.0283	ns	.0971	ns	2.93	.73	.0113	ns	.1145	ns
Variety of materials	1.88	.82	.0142	ns	.0096	ns	1.81	.69	.0128	ns	.1426	.05
Materials effectively support instruction	3.93	.76	.1102	.05 (-)	.0032	ns	4.14	.58	.0512	ns	.0084	ns
Presentation clear	4.12	.87	.0395	ns	.0009	ns	3.98	.70	.0093	ns	.0542	.05
Lesson adapted to different levels	3.31	.98	.1567	.05 (-)	.0060	ns	3.32	.72	.0623	ns	.1792	.01
High degree of students success	3.58	.62	.1455	.01 (-)	.0262	ns	3.79	.52	.0108	ns	.1457	.01

Table 2-continued

Variables	First 3 Weeks $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	Rest of Year $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Content related to students interests, backgrounds	3.41	.87	.2156	.05 (-)	.0000	ns	3.49	.63	.0112	ns	.0470	ns
Reasonable work standards	3.97	.89	.0140	ns	.0013	ns	3.86	.72	.0047	ns	.0218	ns

Table 3

Results for Positive Affect Variables and Class Heterogeneity in English Classes

Variables	First 3 Weeks						Rest of Year					
	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
States desired attitudes	3.53	.97	.0051	ns	.0034	ns	2.86	.83	.1247	.05 (-)	.0000	ns
Amount of positive reinforcement	3.06	1.05	.0358	ns	.0182	ns	2.92	.74	.2922	.001 (-)	.0161	ns
Teacher's listening skills	3.38	.84	.0547	ns	.0102	ns	3.23	.75	.1577	.01 (-)	.0008	ns
Teacher expresses feelings	3.24	.95	.0115	ns	.0004	ns	2.72	.81	.2527	.001 (-)	.0281	ns
Teacher receptive to student input	3.58	.83	.1868	.05 (-)	.0242	ns	3.26	.71	.1699	.01 (-)	.0015	ns
Teacher oriented to student needs	3.44	1.10	.2172	.05 (-)	.0017	ns	3.65	.70	.2142	.001 (-)	.0017	ns
Nurturance of affective skills	2.67	1.14	.1509	.05 (-)	.0031	ns	2.26	.82	.0715	.05 (-)	.0028	ns

Table 3-continued

Variables	First 3 Weeks $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	Rest of Year $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Classroom climate: task-oriented	4.44	.75	.0013	ns	.0001	ns	4.12	.66	.0268	ns	.0123	ns
Teacher encourages group cohesiveness	2.99	.94	.1099	ns	.2530	.01	2.35	.55	.0441	ns	.0361	ns

Table 4

Results for Instructional Management Variables and Class Heterogeneity in English Classes

Variables	First 3 Weeks						Rest of Year					
	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Variety of Materials	1.88	.82	.0142	ns	.0096	ns	1.81	.69	.0128	ns	.1426	.05
Materials ready	4.32	.68	.0408	ns	.0529	ns	4.54	.50	.0210	ns	.0975	ns
Materials effectively support instruction	3.93	.76	.1102	.05 (-)	.0032	ns	4.14	.58	.0512	ns	.0084	ns
Clear directions	4.12	.95	.0151	ns	.0315	ns	4.19	.67	.0093	ns	.0024	ns
Distracting mannerisms	1.42	.95	.0006	ns	.0224	ns	1.33	.57	.0214	ns	.0088	ns
Eye contact	4.17	.83	.0538	ns	.0157	ns	3.87	.85	.0858	ns	.0135	ns
Presentation clear	4.12	.87	.0395	ns	.0009	ns	3.99	.70	.0093	ns	.0542	.05
Provides and seeks rationale, analysis	3.78	.86	.1831	.05 (-)	.0001	ns	3.36	.75	.0418	ns	.0069	ns

Table 5

Results for Behavior Management Variables and Class Heterogeneity in English Classes

Variables	First 3 Weeks $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	Rest of Year $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Amount of positive reinforcement	3.06	1.05	.0358	ns	.0182	ns	2.92	.74	.2922	.001 (-)	.0161	ns
Teacher signals appropriate behavior	3.15	.81	.0040	ns	.0256	ns	3.20	.84	.0532	ns	.0107	ns
Teacher reinforces inattentive behavior	2.13	.74	.0005	ns	.0344	ns	2.30	.65	.0281	ns	.0264	ns
Consistency in dealing with behavior	3.65	1.02	.0002	ns	.0105	ns	3.79	.73	.1059	.01 (-)	.0064	ns
Amount of disruptive behavior	1.82	.87	.0007	ns	.0750	ns	1.39	.43	.0658	ns	.0034	ns
Source of disruptive behavior	1.86	.69	.3181	.01 (+)	.0006	ns	2.18	.72	.0001	ns	.0292	ns
Stops disruptive behavior quickly	3.45	1.21	.0021	ns	.0196	ns	3.22	.85	.0023	ns	.0002	ns
Gives rules and procedures in response to disruptive behavior	2.95	1.14	.0079	ns	.0850	ns	1.71	.76	.0000	ns	.0143	ns

Table 5-continued

Variables	First 3 Weeks $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	Rest of Year $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Criticizes/justifies authority in response to disruptive behavior	1.52	.83	.0332	ns	.0252	ns	1.50	.55	.0159	ns	.0048	ns
Punishes disruptive behavior	1.42	.64	.0157	ns	.0012	ns	1.30	.54	.0688	ns	.0067	ns
Ignores disruptive behavior	2.04	.92	.0503	ns	.0918	ns	1.85	.91	.0469	ns	.0572	ns
Conference in response to disruptive behavior	.16	.23	.0720	ns	.0035	ns	.18	.26	.0942	ns	.0278	ns
Amount of inappropriate behavior	2.26	1.08	.0000	ns	.0002	ns	2.29	.86	.0180	ns	.0006	ns
Inappropriate behavior stopped quickly	3.31	1.51	.0093	ns	.0959	ns	3.51	1.00	.0222	ns	.0179	ns
Gives rules and procedures in response to inappropriate behavior	2.36	1.10	.0468	ns	.0986	ns	1.57	.50	.0137	ns	.0069	ns

Table 5-continued

Variable	First 3 Weeks $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	Rest of Year $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Criticizes/justifies authority in response to inappropriate behavior	1.34	.82	.0133	ns	.0013	ns	1.54	.45	.0984	ns	.0036	ns
Punishes inappropriate behavior	1.55	.90	.0026	ns	.0748	ns	1.19	.30	.0001	ns	.0019	ns
Ignores inappropriate behavior	2.13	1.23	.0001	ns	.0381	ns	2.16	.87	.0339	ns	.0012	ns
Conference in response to inappropriate behavior	.40	.74	.0803	ns	.0137	ns	.52	.42	.0254	ns	.0055	ns
Desists inappropriate behavior	3.24	1.15	.0923	ns	.2291	.05	2.97	.66	.0018	ns	.0268	ns

Table 6

Results for Student Engagement Ratings and Class Heterogeneity in English Classes

Variables	First 3 Weeks $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$	Rest of Year $\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Off-task, unsanctioned	.03	.02	.0446	ns	.0152	ns	.06	.06	.1396	.01 (+)	.0797	.05
Dead time	.03	.04	.1351	ns	.0486	ns	.04	.05	.0357	ns	.0454	ns
On-task, academic	.65	.15	.0593	ns	.0157	ns	.72	.14	.0154	ns	.1743	.05
On-task	.92	.06	.1996	.01 (-)	.0017	ns	.89	.09	.1290	.05 (-)	.0898	.05

Table 7

Teacher Self-Report Variables and Class Heterogeneity in English Classes

Variable	Possible Response Range	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Percent time devoted to individuals in class	0-100	20	13	.1055	ns	.2282	.05
Percent time devoted to subgroups in class	0-100	18	14	.0143	ns	.0274	ns
Percent time devoted to whole class instruction	0-100	62	19	.0954	ns	.0406	ns
Frequency of use of ability grouping	1-3	2.33	.62	.1162	.05 (+)	.0233	ns
Frequency of use of individualized work, contracts	1-3	2.07	.78	.0446	ns	.0251	ns
Frequency of use of different levels of materials	1-3	2.18	.79	.0257	ns	.0056	ns
Frequency of use of special attention, conferences to cope with varying ability levels in class	1-3	2.22	.64	.0948	ns	.0019	ns
Frequency of use of peer tutoring	1-3	2.15	.72	.3440	.01 (+)	.0102	ns

Table 7-continued

Variable	Possible Response Range	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Frequency of use of differential grading and testing to cope with varying ability levels in class	1-3	2.37	.84	.0998	ns	.0238	ns
Frequency of teaching to higher or middle group (only) to cope with varying ability levels in class	1-3	1.96	.76	.0103	ns	.0935	ns
Frequency of use of help from student teacher, aides, or resource teachers to cope with varying ability levels in class	1-3	1.63	.69	.1145	ns	.1145	ns
Emphasis on effort in grading middle or mixed ability classes	1-5	3.89	1.22	.0212	ns	.0685	ns
Extent to which academic range of students considered a problem or constraint	1-5	2.52	.89	.0552	ns	.1150	ns
Extent of teacher concern with meeting needs of different kinds of students	1-5	4.52	.64	.0369	ns	.0356	ns

C2

C3

Table 7-continued

Variable	Possible Response Range	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Extent of teacher concern with whether each student gets what he needs	1-5	4.33	.68	.0215	ns	.0180	ns

Table 8

Class Mean End-of-year Outcome Variables

and Class Heterogeneity in English Classes

	$\bar{x}$	SD	Difference in R^2 Main Effect	$p \leq$	Difference in R^2 Interaction	$p \leq$
Achievement gains	76.12	4.24	.0064	ns	.0007	
Student Ratings of the Teacher	61.34	3.98	.0001	ns	.0347	ns

Table 9

Learning Gains of High- and Low-ability Students

Within English Classes

Class Number	Sigma	Percent of Students with Expected[1] or Above Expected Achievement Gains	
		Low-ability[2]	High-ability[3]
3.5	14.45	75	50
3.6	11.29	78	0
5.6	10.19	100	50
6.5	13.89	100	50
7.4	14.21	50	86
7.5	12.35	100	86
11.3	13.41	100	72
11.5	14.04	40	70
14.1	17.2[illegible]	63	86
14.2	14.7[illegible]	80	60
16.1	17.[illegible]	33	89
16.3	15.41	67	86
18.2	13.92	50	66
20.5	9.12	29	66
22.4	9.53	67	50
22.6	9.95	64	100
25.5	6.59	78	nd
26.3	14.91	33	50
27.1	11.7	100	100
27.4	11.35	100	50

C7

Table 9-continued

Class Number	Sigma	Percent of Students with Expected[1] or Above Expected Achievement Gains: Low-ability[2]	High-ability[3]
38.4	4.62	100	nd
45.5	10.59	100	91
46.1	8.75	71	80
48.2	9.03	100	43
48.5	8.33	100	67
53.3	8.28	33	71
53.6	9.37	100	63
Means	11.65	74.48	67.28
r for Correlation with sigma		-.33	.17
$p \leq$		.05	ns

[1]Expected Gains = -.25 to +.25 residual achievement gain.

[2]Low-ability = Students scoring in the bottom 33rd percentile on entering achievement test.

[3]High-ability = Students scoring in the top 33rd percentile on entering achievement test.

Table 10

Descriptive Data for Five Very Heterogeneous English Classes

Class Number	Teacher Management Factor Score	Mean Entering CAT	Grade Level	Sigma of CAT Scores	Grade Level Equivalent Range	Number of Students	Mean Achievement Residual Gain	Mean SRT
A (7.4)	114	39.5	7	14.21	1.9-11.4	20	.311	64.4
B (11.5)	113	47.7	8	14.04	2.4-11.4	24	-.021	60.6
C (14.1)	86	37.8	7	17.24	2.4-10.9	22	-.018	57.8
D (16.')	118	44.8	7	17.17	2.4-11.6	22	-.252	63.2
E (26.3)	108	40.3	7	14.91	2.6-11.4	17	-.265	67.6
Mean for all middle-ability English classes	98.5	43.0		11.65			-.015	61.3

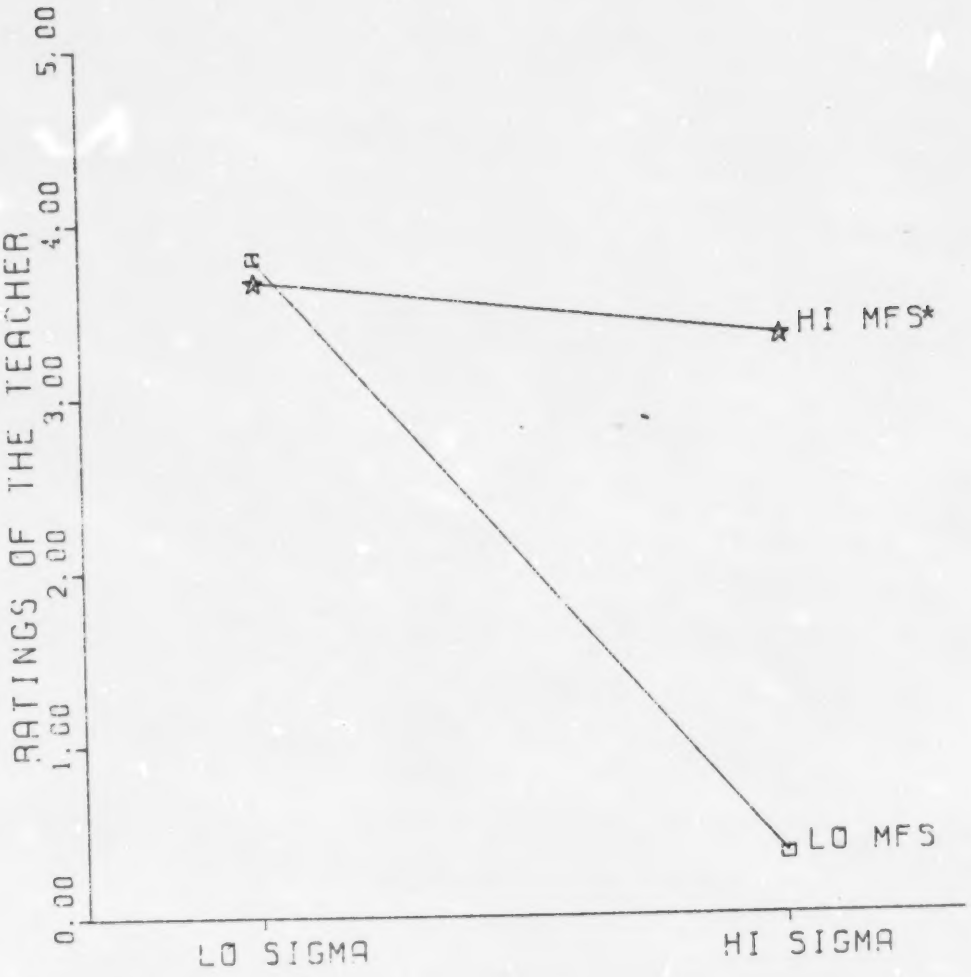

Figure 1. English classes, rest of year: Component Rating 2: Students' attention spans are considered. (*MFS is Classroom Management Factor Score.)

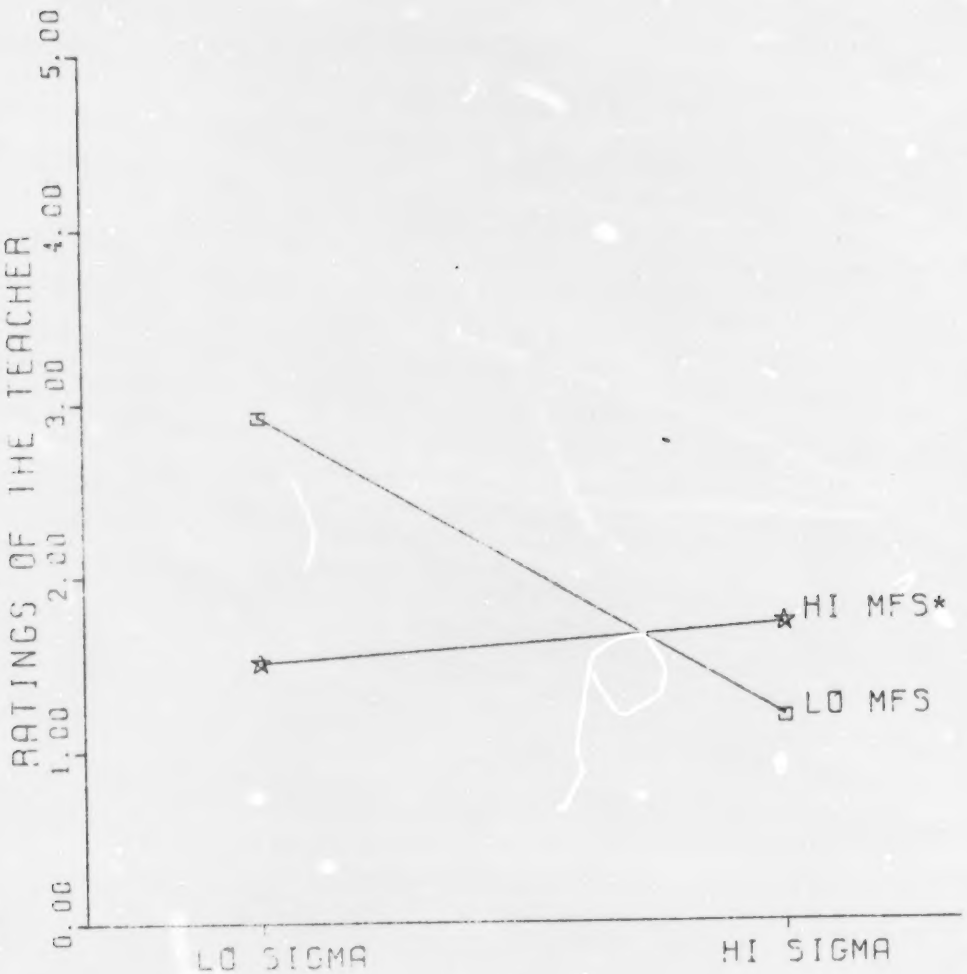

Figure 2. English classes, rest of year: Component Rating 5: Variety of materials. (*MFS is Classroom Management Factor Score.)

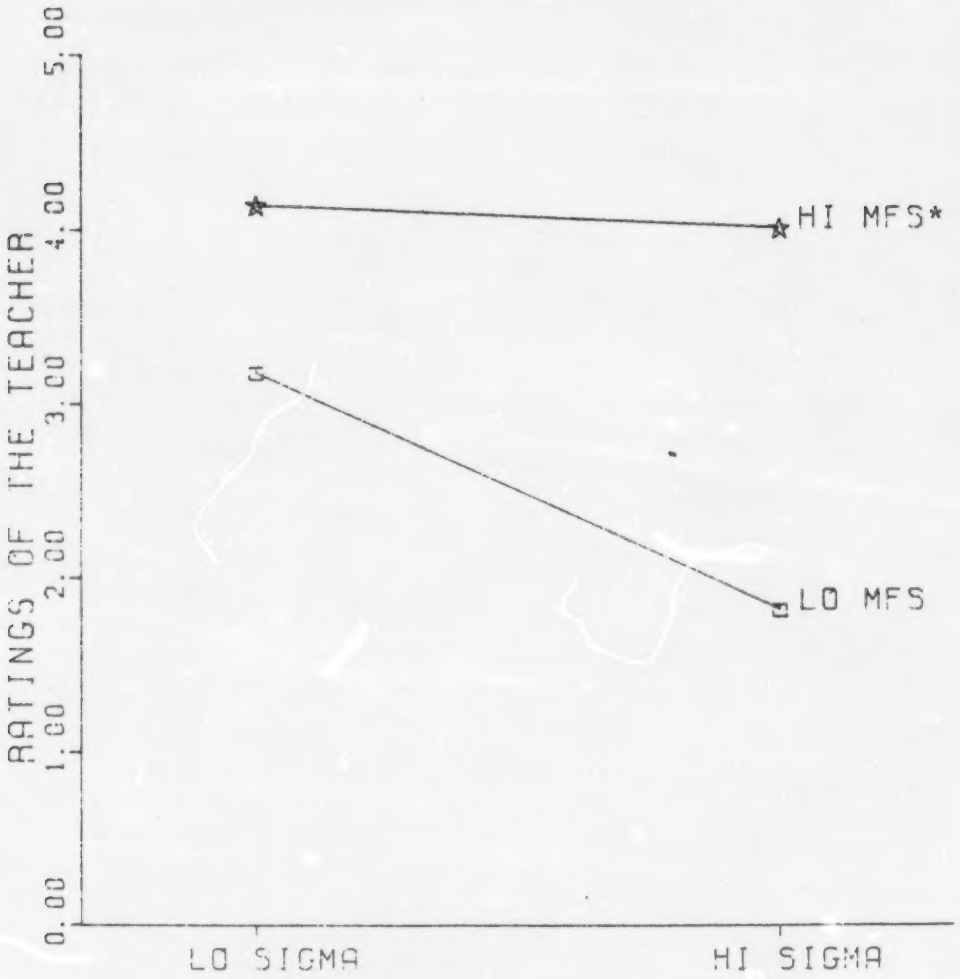

Figure 3. English classes, rest of year: Component Rating 11: Teacher's presentation of lesson is clear. (*MFS is Classroom Management Factor Score.)

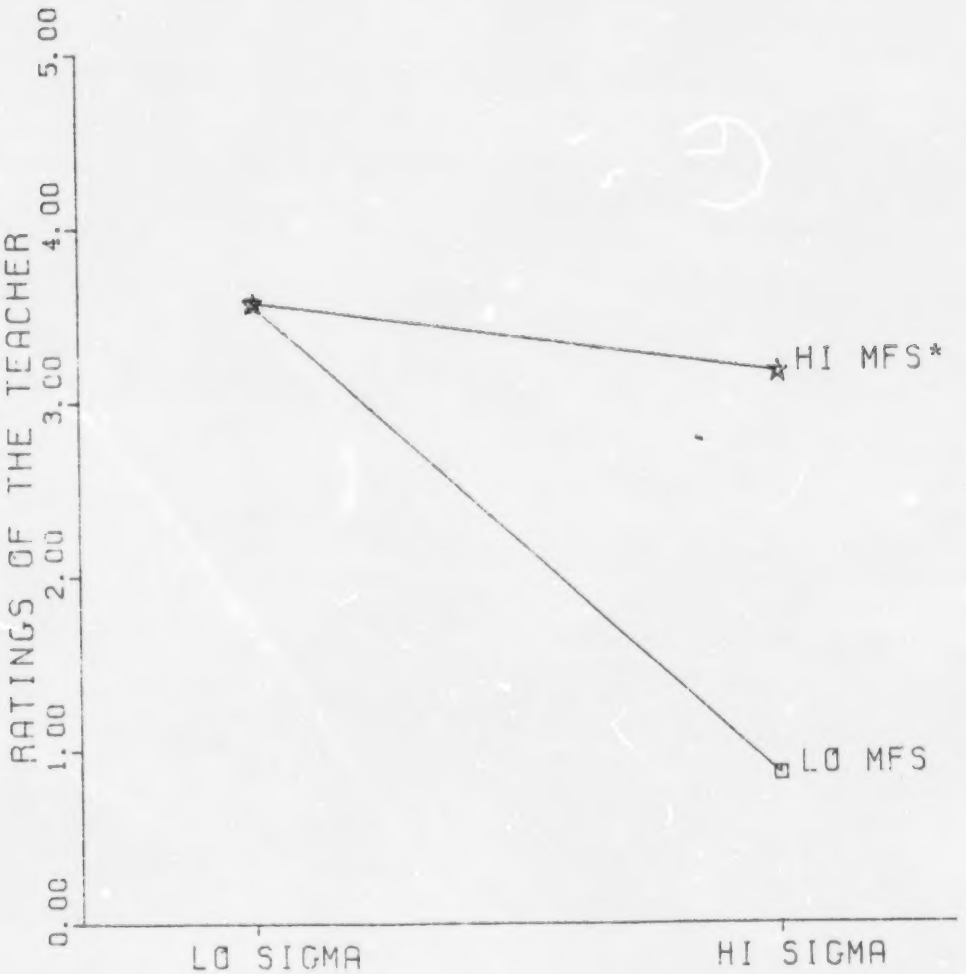

Figure 4. English classes, rest of year: Component Rating 12: Lesson is adapted to different academic levels of students. (*MFS is Classroom Management Factor Score.)

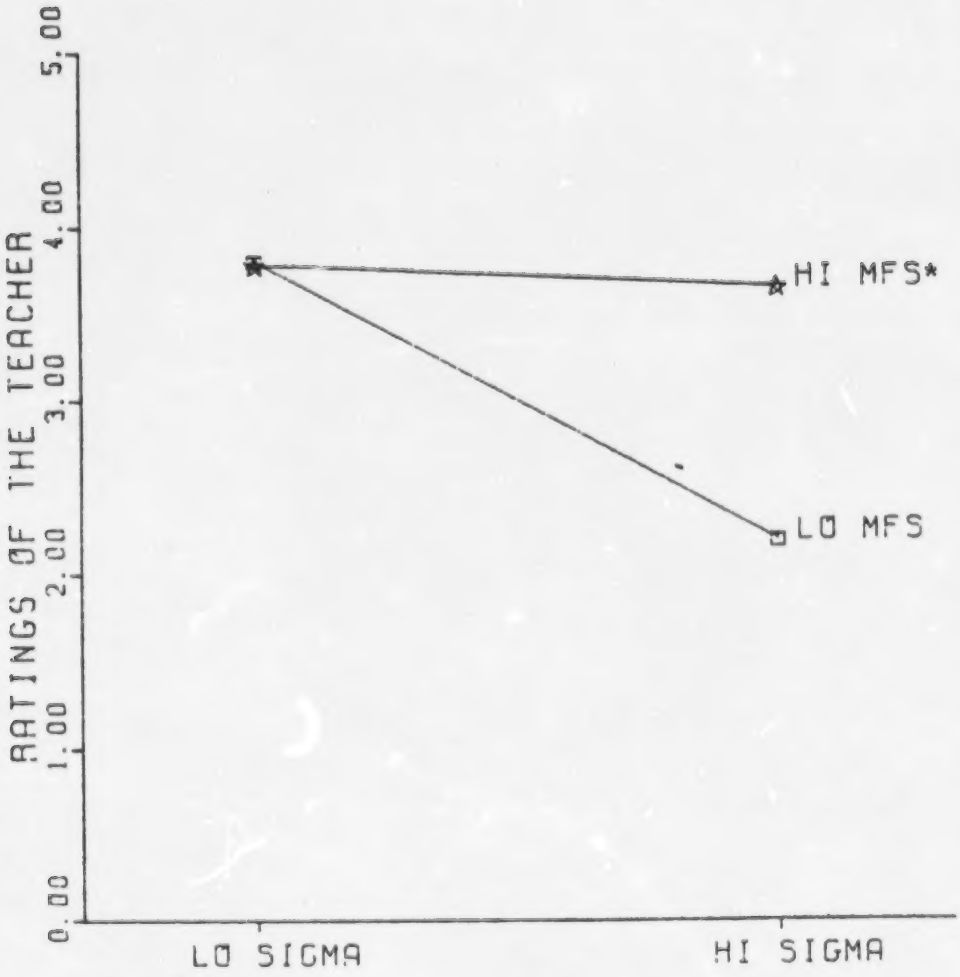

Figure 5. English classes, rest of year: Component Rating 15: High degree of student success. (*MFS is Classroom Management Factor Score.)

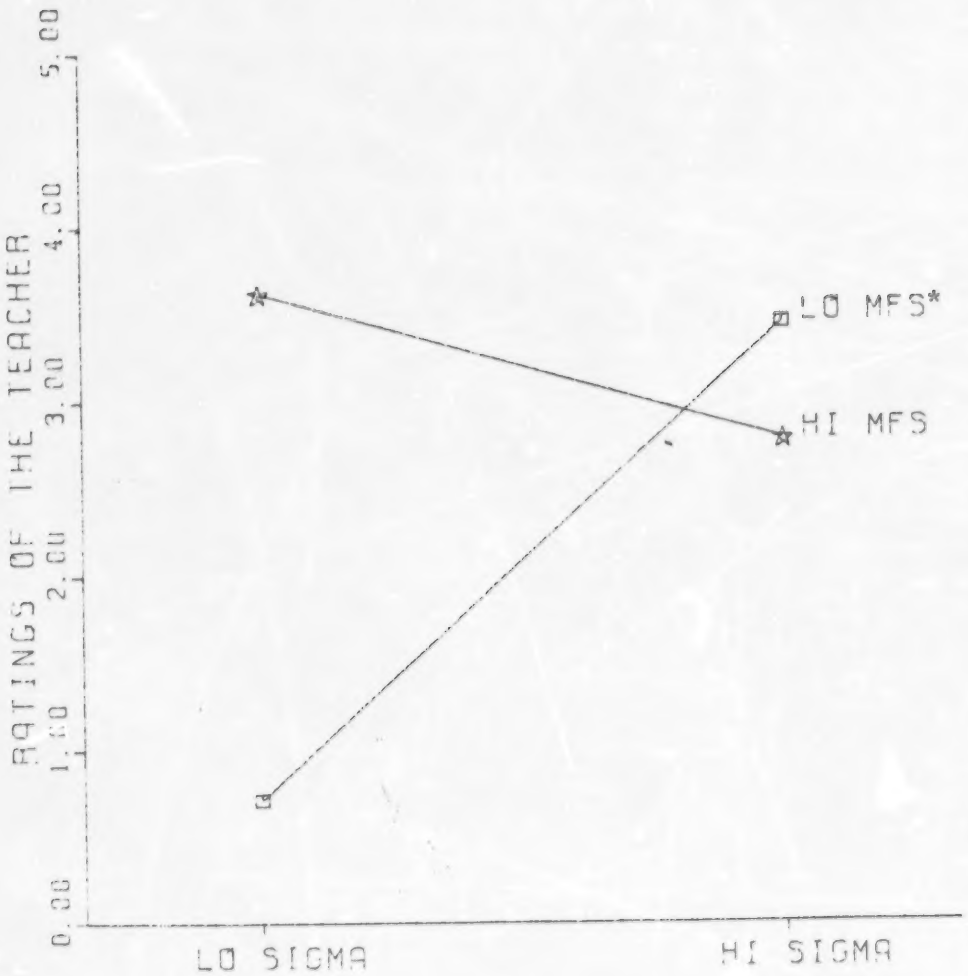

Figure 6. English classes, first three weeks: Component Rating 36: Teacher encourages group cohesiveness. (*MFS is Classroom Management Factor Score.)

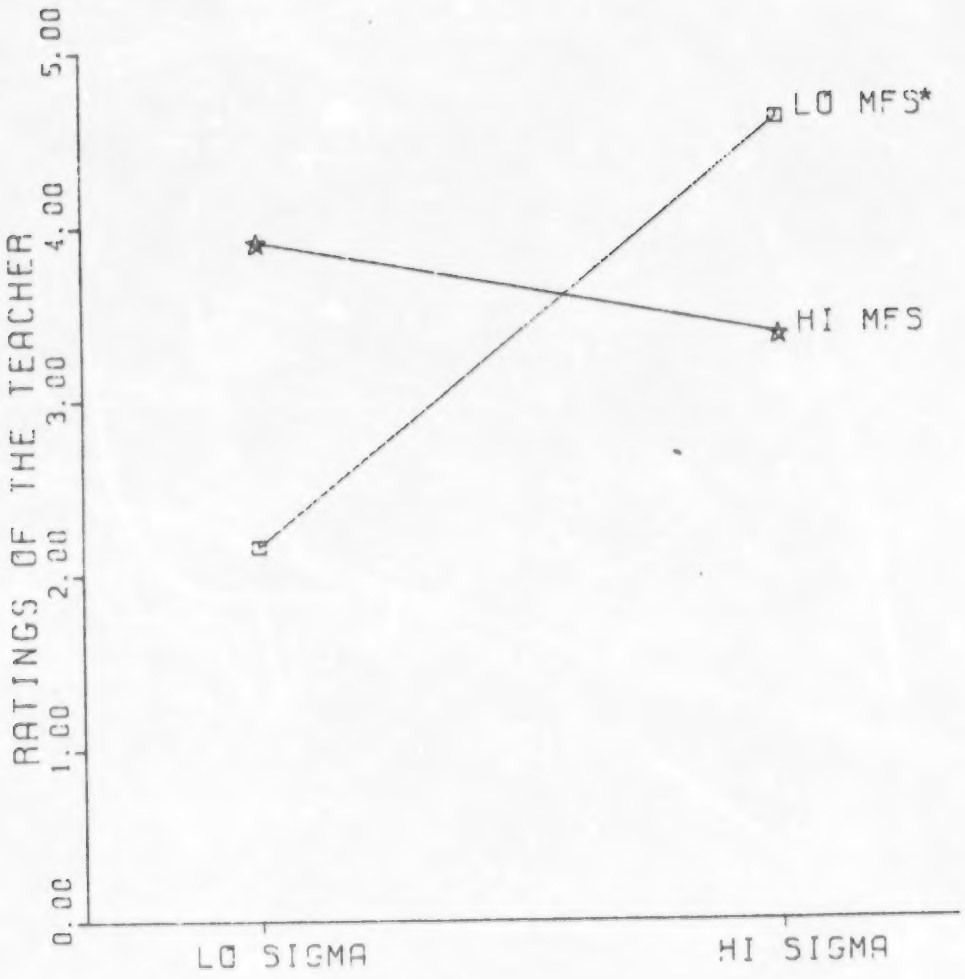

Figure 7. English classes, first three weeks: Component Rating 44: Teacher desists inappropriate behavior. (*MFS is Classroom Management Factor Score.)

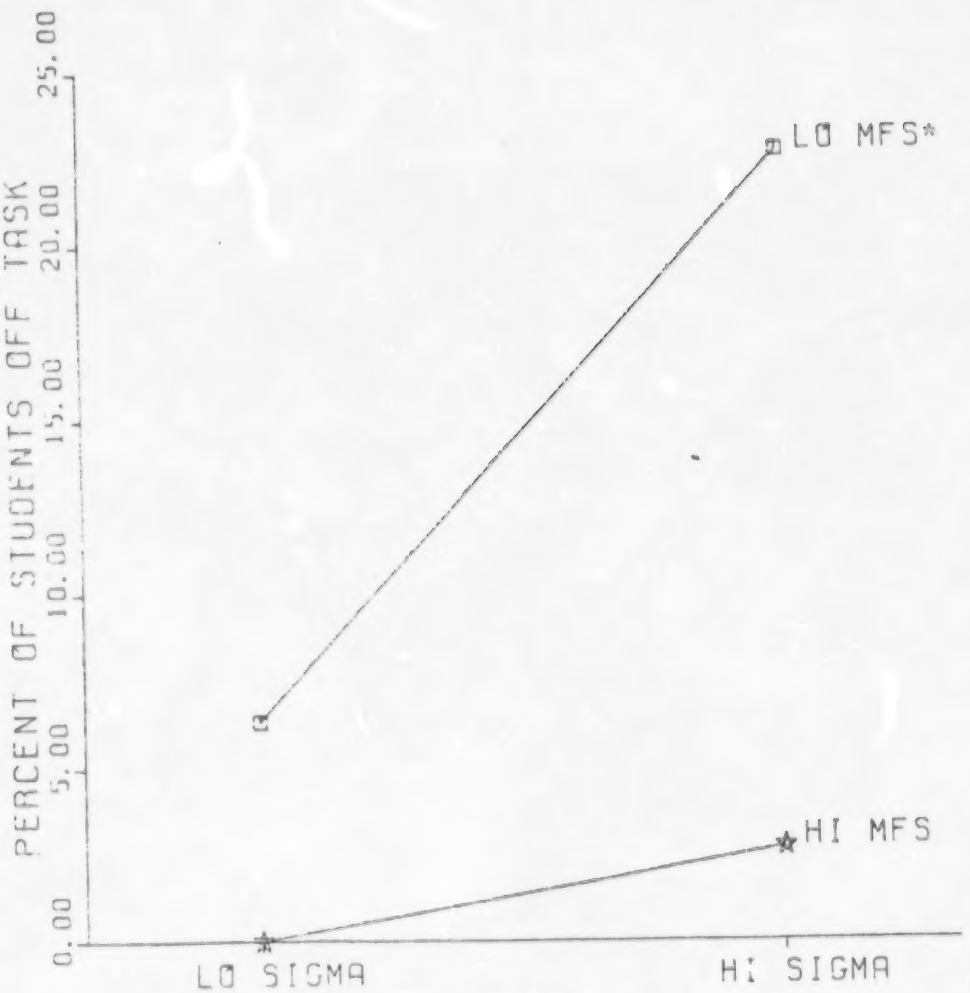

Figure 8. English classes, rest of year: Percentage of students off-task, unsanctioned. (*MFS is Classroom Management Factor Score.)

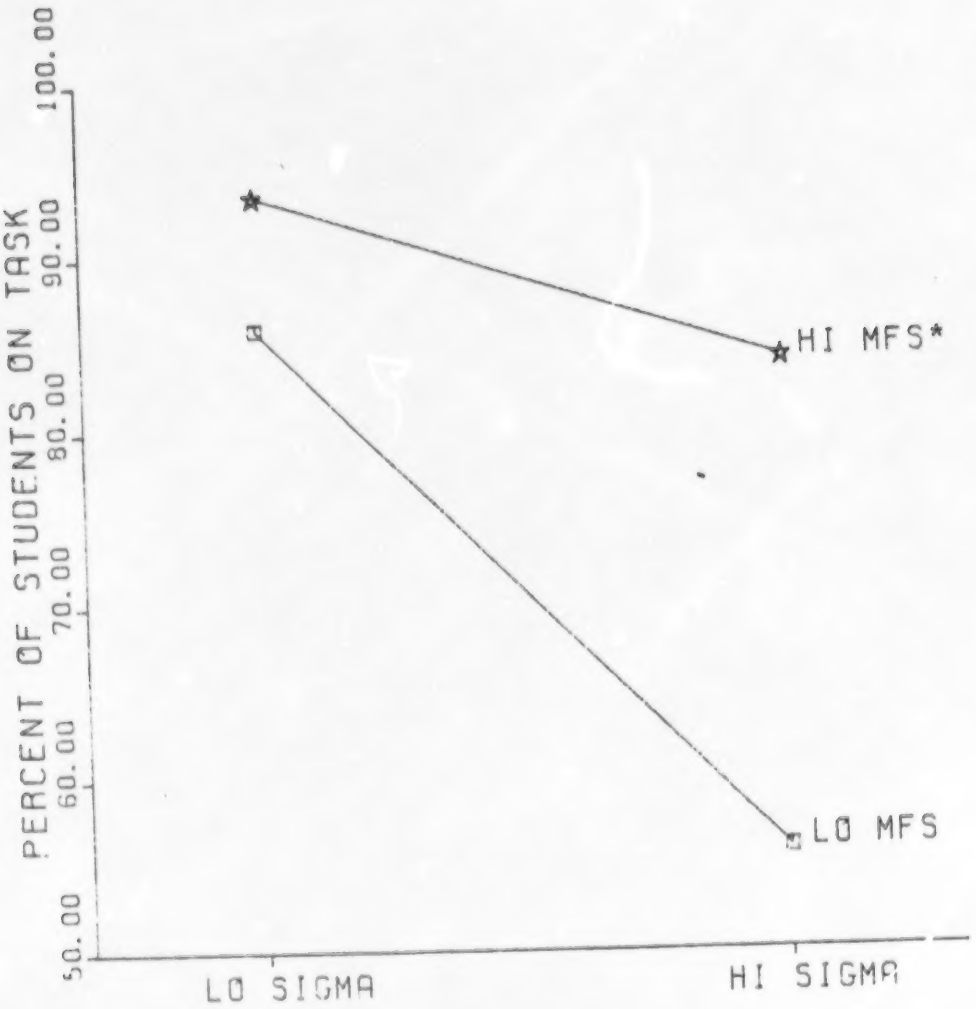

Figure 9. English classes, rest of year: Percentage of students on-task. (*MFS is Classroom Management Factor Score.)

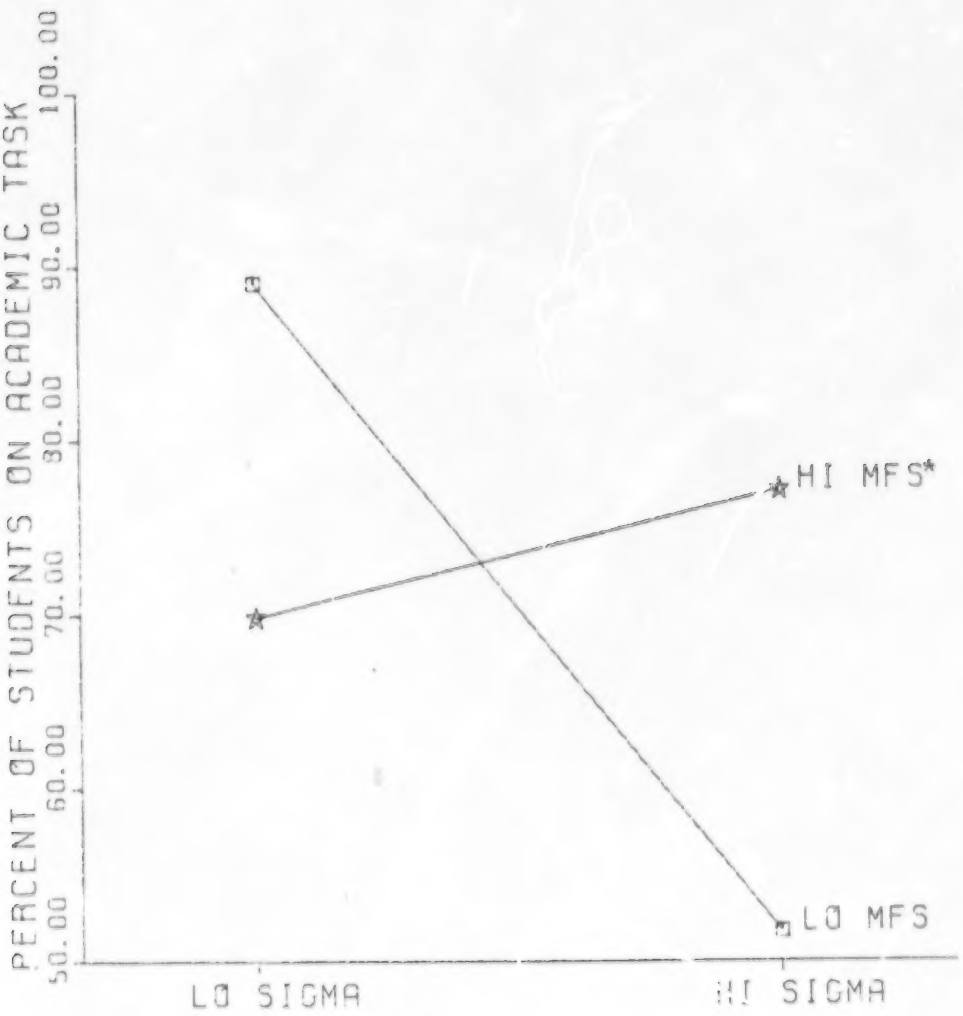

Figure 10. English classes, rest of year: Percentage of students on-task, academic. (*MFS is Classroom Management Factor Score.)

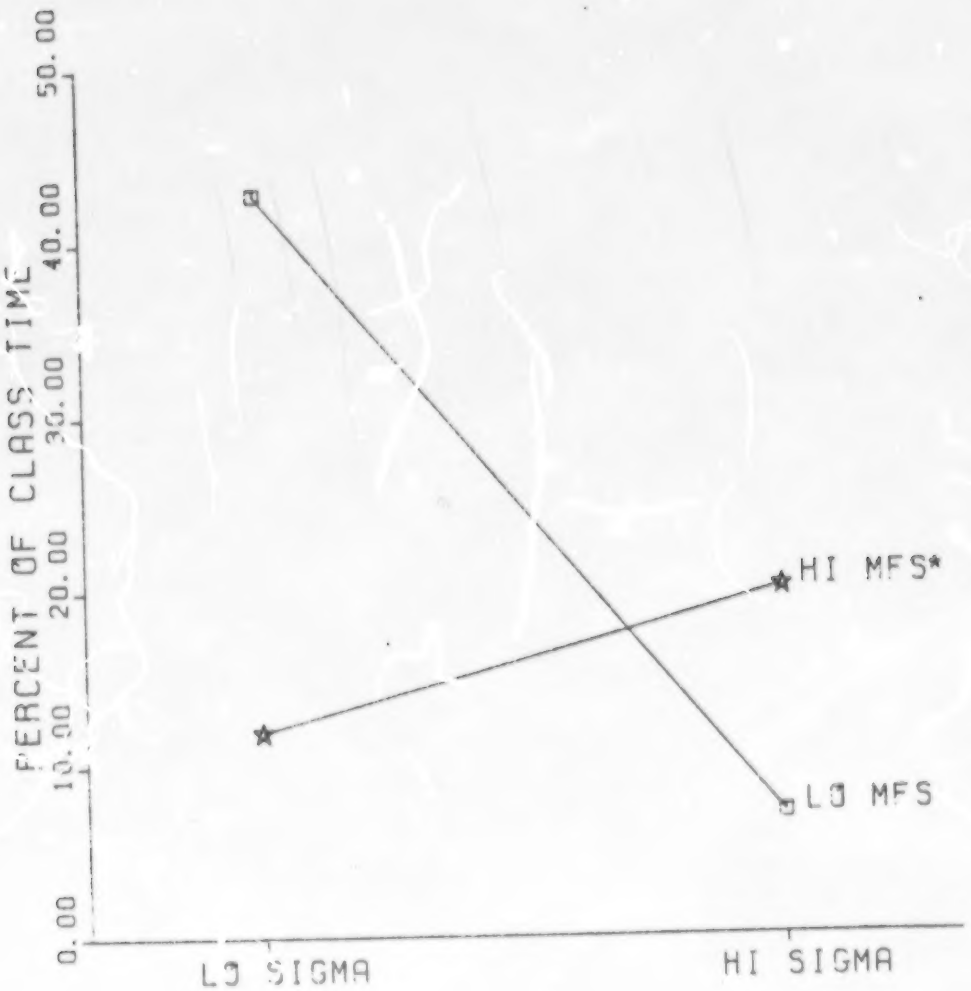

Figure 11. English teacher questionnaire: Teacher estimates of class time devoted to individuals. (*MFS is Classroom Management Factor Score.)